

FUNDAÇÃO UNIVERSIDADE FEDERAL DE RONDÔNIA
DEPARTAMENTO DE MATEMÁTICA E ESTATÍSTICA
BACHARELADO EM ESTATÍSTICA

ROBSON PRATA FROTA

**APLICAÇÃO DA METODOLOGIA BOX E JENKINS PARA PREVISÃO
DA PRECIPITAÇÃO PLUVIOMÉTRICA NO ESTADO DE RONDÔNIA**

Jl-PARANÁ

2018

ROBSON PRATA FROTA

**APLICAÇÃO DA METODOLOGIA BOX E JENKINS PARA PREVISÃO
DA PRECIPITAÇÃO PLUVIOMÉTRICA NO ESTADO DE RONDÔNIA**

Trabalho de Conclusão de Curso submetido ao Departamento de Matemática e Estatística, da Fundação Universidade Federal de Rondônia, *Campus* de Ji-Paraná, como parte dos requisitos para obtenção do título de Bacharel em Estatística.

Orientador: Fernando Cardoso

JÍ-PARANÁ

2018

Dados Internacionais de Catalogação na Publicação
Fundação Universidade Federal de Rondônia
Gerada automaticamente mediante informações fornecidas pelo(a) autor(a)

F941a Frota, Robson Prata.

Aplicação da metodologia Box e Jenkins para revisão da precipitação pluviométrica no Estado de Rondônia / Robson Prata Frota. -- Ji-Paraná, RO, 2018.

44 f. : il.

Orientador(a): Prof.^a Ma. Fernando Cardoso

Trabalho de Conclusão de Curso (Graduação em Estatística) - Fundação Universidade Federal de Rondônia

1.Precipitação. 2.Previsão. 3.Box e Jenkins. I. Cardoso, Fernando. II. Título.

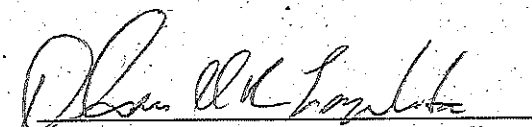
CDU 519.246.8



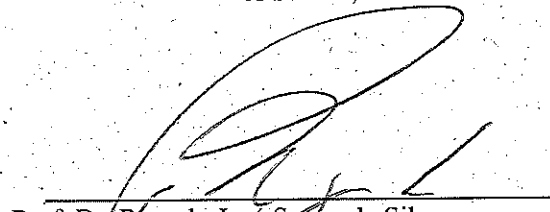
**ATA DA SESSÃO DE APRESENTAÇÃO E DEFESA DE
TRABALHO DE CONCLUSÃO DE CURSO
BACHARELADO EM ESTATÍSTICA**

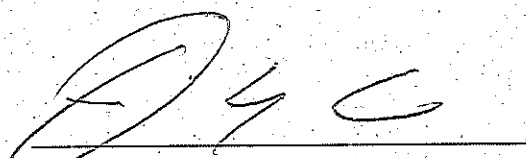
Aos cinco dias do mês de julho de 2018, realizou-se na Sala do Laboratório de Estatística, no *Campus* de Ji-Paraná, a Sessão de Apresentação Trabalho de Conclusão de Curso (TCC), do Aluno Robson Prata Frota, com o título: “**APLICAÇÃO DA METODOLOGIA BOX E JENKINS PARA PREVISÃO DA PRECIPITAÇÃO PLUVIOMÉTRICA NO ESTADO DE RONDÔNIA**”. Os trabalhos foram iniciados às 9 horas pelo presidente da Banca Examinadora, constituída por: Prof. Fernando Luiz Cardoso (Orientador), Prof. Dr. Dilson Henrique Ramos Evangelista (Membro), e Prof. Dr. Ricardo José Souza da Silva (Membro). A Banca Examinadora, tendo decidido aceitar o Trabalho, passou à arguição pública do acadêmico. Encerrando os trabalhos de arguição às 10h10min a Banca deu parecer final de **APROVADO**, com a nota 80 (oitenta) resultado da média aritmética das notas individuais atribuídas pelos membros da Banca Examinadora. Proclamado o resultado final pelo presidente da Banca, foram encerrados os trabalhos. Para constar, lavrou-se a presente Ata, que é assinada pelos membros da Banca Examinadora e pelo acadêmico.

Ji-Paraná, 05 de Julho de 2018.

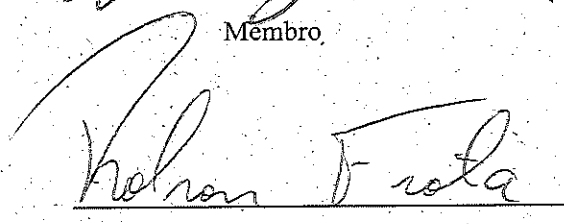

Prof. Dr. Dilson Henrique R. Evangelista

Membro


Prof. Dr. Ricardo José Souza da Silva
Membro


Prof. Ms. Fernando L. Cardoso

Orientador e Presidente da Banca


ROBSON PRATA FROTA

Acadêmico

A Deus

A minha esposa

Aos meus pais e irmãos

Aos amigos que me deram força

DEDICO

AGRADECIMENTOS

Agradeço primeiramente, a Deus por me dar saúde e muita força para superar todas as dificuldades.

A universidade UNIR e todo o corpo docente por ter nos dado a oportunidade de realizar este curso.

Ao meu orientador, pela força, suporte, apoio e incentivos, durante o processo de realização desse trabalho.

A minha amada esposa Meline, que com muito carinho e apoio, não mediu esforços para que eu chegasse até esta etapa de minha vida.

Aos meus pais, em especial a minha mãe Nelita por todo amor, por todo incentivo, por nunca me deixar desistir, por ser o maior exemplo de superação que tenho na vida.

Ao meu irmãos e amigos que me deram força.

Enfim, agradeço a todos que direta ou indiretamente fizeram parte da minha formação.

“Ninguém é tão grande que não possa aprender, nem tão pequeno que não possa ensinar.”

Esopo, Escritor (620 a.C. - 564 a.C.)

RESUMO

As mudanças climáticas têm chamado a atenção de muitos pesquisadores nos últimos anos. Entender como os elementos climáticos funcionam, e poder prever qual o seu comportamento no futuro, teria grande influência na economia, tanto podendo evitar tragédias como auxiliando no manejo da terra, potencializando o cultivo de várias culturas. É de grande importância para a agricultura, pecuária, piscicultura e outros, entender o comportamento do clima em uma determinada região. Este trabalho realizou a análise de séries temporais nos dois extremos do estado de Rondônia, nas cidades de Porto Velho e Vilhena, entender o comportamento do clima nessas cidades é de grande importância para a economia do Estado. foi utilizado a metodologia Box e Jenkins para fazer previsões da precipitação pluviométrica, nas cidades de Porto Velho e Vilhena, a modelagem Box e Jenkins gerou quatro modelos SARIMA, um para cada elemento em estudo, a análise de previsão foi bem-sucedida para todos os modelos, a previsão apresentou um erro pequeno e um comportamento semelhante ao das series de dados originais.

Palavras-chave: precipitação; previsão: Box e Jenkins.

ABSTRACT

The climate changes has attracted the attention of many researchers in recent years. Understanding how climate elements work, and being able to predict how they will behave in the future, would have a major influence on the economy, both avoiding tragedies and aiding land management, potentializing the cultivation of various crops. Is of great importance for agriculture, livestock, fish farming and others, understanding the behavior of the climate in a given region. This work carried out the analysis of time series in the two extremes of the state of Rondônia, the cities of Porto Velho and Vilhena are in opposite parts of the state, one to the north and another to the south, to understand the behavior of the climate in these cities is of great importance for the economy of the state. The Box & Jenkins methodology was used to do make rainfall forecasts in the cities of Porto Velho and Vilhena, the Box & Jenkins model generated four SARIMA models, one for each element under study, the prediction analysis was successful for all models, the forecast presented a small error and a behavior similar to that of the original data series.

Key-words: precipitation; predicted; Box & Jenkins.

LISTA DE ILUSTRAÇÕES

Figura 1. Localizações das cidades de Porto Velho e cidade de Vilhena no estado de Rondônia. (Fonte: WIKIPÉDIA, 2017.)	14
Figura 2. Um processo estocástico interpretado como uma família de variáveis aleatórias (MORETTIN e TOLOI, 2014, p. 20).	16
Figura 3. Um processo estocástico interpretado como uma família de trajetórias (MORETTIN e TOLOI, 2014, p. 20).	17
Figura 4. Ciclo iterativo de Box e Jenkins.	21
Figura 5. Representação gráfica da série original referente a precipitação pluviométrica da cidade de Porto Velho RO, no período de Julho de 2007 até Agosto de 2015.....	26
Figura 6. Representação gráfica da série original referente a precipitação pluviométrica da cidade de Vilhena RO, no período de Agosto de 2008 até Agosto de 2015.	26
Figura 7. Gráfico BoxPlot apresentando os dados em medias mensais das séries originais da precipitação pluviométrica da cidade de Porto Velho RO.	27
Figura 8. Gráfico BoxPlot apresentando os dados em medias mensais das séries originais da precipitação pluviométrica da cidade de Vilhena RO.	28
Figura 9. Gráfico da decomposição das series da precipitação pluviométrica da cidade de Porto Velho.	29
Figura 10. Gráfico da decomposição das series da precipitação pluviométrica da cidade de Vilhena.	29
Figura 11. Representação gráfica das funções de autocorrelação (ACF) e autocorrelação parcial (PACF) da série diferenciada sazonalmente da precipitação pluviométrica das cidades de Porto Velho Vilhena.....	31
Figura 12. Gráfico da previsão da série de precipitação pluviométrica da cidade de Porto Velho RO, no período de Setembro de 2015 a Maio de 2016.	34
Figura 13. Gráfico da previsão da série de precipitação da cidade de Vilhena RO, no período de Setembro de 2015 a Maio de 2016.	35

LISTA DE TABELAS

Tabela 1. Resultados dos testes de sazonalidade e tendência para a precipitação pluviométricas das cidades de Porto Velho e Vilhena.	30
Tabela 2. Resultados dos testes Shapiro-Wilk, Kolmogorov-Smirnoff, Box-Pierce e Ljung-Box para os resíduos dos modelos escolhidos.	32
Tabela 3. Estatísticas do modelo SARIMA (p,d,q) x (P,D,Q) _s , da precipitação pluviométrica da cidade de Porto Velho.	33
Tabela 4. Estatísticas do modelo SARIMA (p,d,q) x (P,D,Q) _s , da precipitação pluviométrica da cidade de Vilhena.	33
Tabela 5. Valores reais e valores previstos do modelo SARIMA(1,0,2) x (1,0,1) ₁₂ , no período de Setembro de 2015 a Maio de 2016, da precipitação pluviométrica da cidade de Porto Velho.	36
Tabela 6. Valores reais e valores previstos do modelo SARIMA(1,0,1) x (1,0,1) ₁₂ , no período de Setembro de 2015 a Maio de 2016, da precipitação pluviométrica da cidade de Vilhena.	37

LISTA DE ABREVIATURAS

AIC	Critério de Informação de Akaike
ACF	Função de Autocorrelação
ADF	Dickey e Fuller Aumentado
ARFIMA	Autorregressivos Memória Longa
ARMA	Autorregressivos e de Medias Moveis
ARIMA	Autorregressivos integrados e de Medias Moveis
INMET	Instituto Nacional de Meteorologia
MAPE	Erro Percentual Absoluto Médio
PACF	Função de Autocorrelação Parcial
PVH	Porto Velho
SARIMA	Autorregressivo Integrado de Médias Móveis Sazonal
VHA	Vilhena

SUMÁRIO

1	INTRODUÇÃO	12
2	MATERIAIS E MÉTODOS.....	14
	2.1 BASE DE ESTUDO	14
	2.2 METODOLOGIA BOX E JENKINS.....	15
3	RESULTADOS E DISCUSSÕES	26
4	CONCLUSÕES	38
5	REFERÊNCIAS BIBLIOGRÁFICAS	39
6	APÊNDICE A - ROTINA DA ANÁLISE DE SÉRIES TEMPORAIS	41

1 INTRODUÇÃO

O estado de Rondônia segue com um grande crescimento na área agrícola, com o cultivo de arroz, banana, café, mandioca e soja (CARVALHO et al., 2016). A piscicultura é outra fonte de renda que vem crescendo em Rondônia.

Alguns elementos climáticos como temperatura do ar, radiação, velocidade do vento, umidade relativa, precipitação pluviométrica e a evaporação, influenciam diretamente em todos os cultivos citados, estudar e entender como alguns desses elementos podem sofrer alterações pode potencializar o crescimento dessas culturas.

O estado de Rondônia teve um incentivo muito grande na época de sua colonização para o desmatamento e cultivo de diversas culturas agrícola, assim como a pecuária e outros.

Von Randow et al. (2012) destacam que diversos estudos indicam que a substituição da vegetação nativa por pastagem ou culturas agrícolas (como a soja), afeta o clima e ciclo hidrológico através de mudanças no albedo, temperatura da superfície, umidade do solo, regime de evapotranspiração e escoamento.

Na agricultura brasileira a precipitação pluvial é considerada o fator de risco mais importante, pois a chuva excessiva ou a falta dela responde pela maioria dos sinistros agrícolas, (BLAIN, 2009).

A capacidade de prever possíveis mudanças em elementos climáticos como a temperatura do ar, precipitação e evaporação, de um ano para o outro podendo antecipar como eles podem sofrer variações, possibilita o melhor preparo da terra para se adaptar a diversas culturas agrícolas, podendo escolher o melhor cultivo baseando-se nas possíveis mudanças no clima na região pesquisada. (PEREIRA et al., 2015; SILVA et al., 2008.).

O tema das mudanças climáticas e seus impactos têm sido uma pesquisa chave nos últimos anos e têm atraído muita atenção dos pesquisadores em todo o mundo. A mudança climática pode ser definida como qualquer mudança sistemática na evolução em longo prazo no sistema climático que é sustentado por várias décadas ou mais (TABARI e TALAEE 2014).

Estudar e compreender o clima em uma região determinada. é de grande importância para a agricultura e outros setores da economia, assim como o setor ambiental, setor agropecuário, piscicultura entre outros. Utilizar as técnicas de análise de series temporais para

identificar como alguns componentes interferem no comportamento dos elementos climáticos resulta em um meio de entender tais relações existentes nos dados climatológicos, (CARVALHO et al.,2016).

Uma serie temporal e um conjunto de observações ordenadas no tempo de qualquer fenômeno aleatório. A análise de séries temporais consiste em encontrar relações de dependência temporal nos dados, descrevendo seu comportamento para assim poder fazer previsões eficiente nos dados. (PINTO et al., 2015)

Pereira et al. (2015) afirmam que a metodologia Box e Jenkins se mostra bastante eficiente para determinar as correlações entre as séries temporais de diversos elementos climáticos, mesmo com as imprevisibilidades do tempo e do clima com o tratamento correto dos dados a análise se mostra eficiente.

Diante do exposto, este estudo tem como objetivo identificar padrões de comportamento, para fazer previsões eficientes da Precipitação Pluviométrica, em duas cidades do estado de Rondônia, sendo elas Porto Velho e Vilhena, usando a metodologia Box e Jenkins, assim identificando os aspectos sazonal e de tendência para cada elemento citado.

2 MATERIAIS E MÉTODOS

2.1 BASE DE ESTUDO

Os dados foram fornecidos pelo Instituto Nacional de Meteorologia (INMET), o INMET disponibiliza gratuitamente os dados pela internet, através do site (<http://www.inmet.gov.br/portal/index.php?r=estacoes/estacoesautomaticas>).

O sistema de informação meteorológica automática de superfície do INMET trabalha com sub-sistema de coleta de dados, através de sensores que medem as variáveis ambientais; sub-sistema de controle e armazenamento local em *data-logger*; sub-sistema de energia; sub-sistema de comunicações; sub-sistema de banco de dados; e sub-sistema de disseminação de dados aos usuários. (INMET, 2011)

A estação meteorológica coleta dados de minuto a minuto, as informações meteorológicas (temperatura, umidade, pressão atmosférica, precipitação, direção e velocidade dos ventos, radiação solar) representativas da área em que está localizada. A cada hora, estes dados são integralizados e disponibilizados para serem transmitidos, via satélite ou telefonia celular, para a sede do INMET, em Brasília. (OLIVEIRA, 2017)

Para esse trabalho foram utilizados dados de duas localidades do estado de Rondônia, a estação de Porto Velho localizada no Norte do Estado, e a estação de Vilhena localizado na região Sul do estado de Rondônia, a figura 1 traz a localização no mapa.

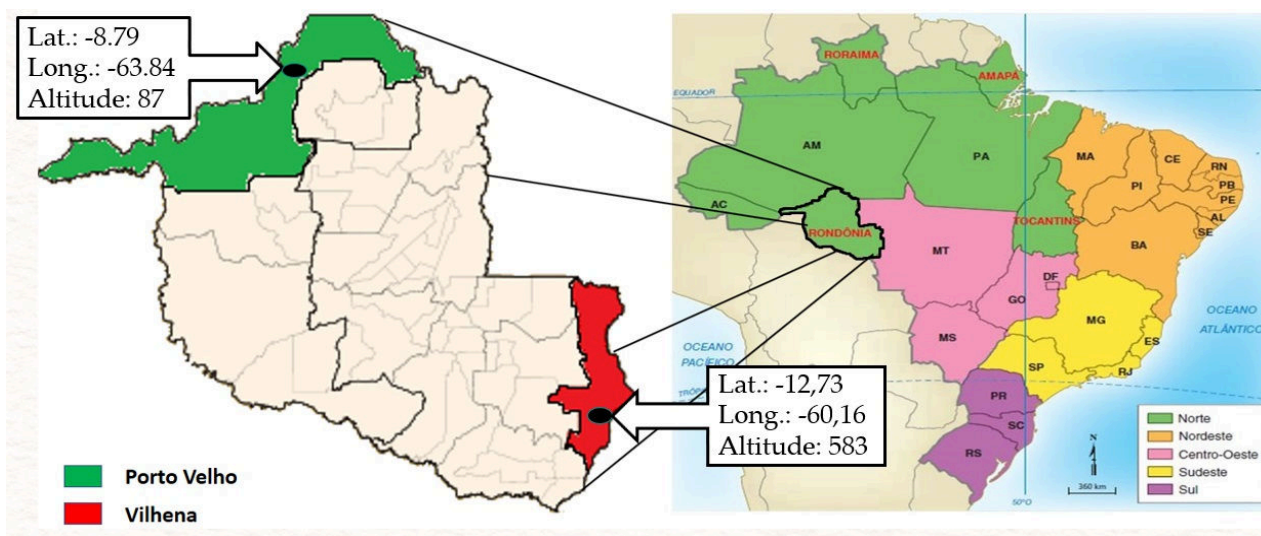


Figura 1. Localizações das cidades de Porto Velho e cidade de Vilhena no estado de Rondônia. (Fonte: WIKIPÉDIA, 2017.)

A estação de Porto Velho foi instalada no dia 11/07/2007, e está localizada na latitude -8.79 e na longitude -63.85, a estação se encontra a 87 metros acima do nível do mar. A estação de Vilhena foi instalada no dia 26/08/2008, e está localizada na latitude -12.73 e longitude -60.16, a estação encontra-se a 583 metros acima do nível do mar. (INMET, 2011)

Será analisado Precipitação Pluviométrica, as análises dos dados foram feitas por observações mensais, foi usado a soma total por mês para a análise da precipitação. Para falhas nos dados foi feita a média das cinco últimas observações e a média das cinco primeiras observações depois foi feita a média das médias.

Os dados utilizados da cidade de Porto Velho são referentes ao período de Julho de 2007 até Maio de 2016 totalizando 108 observações, mas os últimos nove meses serão deixados de fora de todas as análises e utilizados somente para a comparação da previsão desses mesmos meses. Para a cidade de Vilhena o período utilizado é de Agosto de 2008 até Maio de 2016 totalizando 95 observações, sendo deixado de fora os últimos nove meses para comparativo da previsão.

2.2 METODOLOGIA BOX E JENKINS

No próximo item faremos uma breve apresentação sobre séries temporais utilizando como base Morettin e Toloi (2014).

Uma série temporal é um conjunto de observações feitas e armazenadas ao longo do tempo, essas observações devem possuir dependência serial. A série temporal é definida como conjunto de observações quantitativa. Sendo dita como série temporal contínua aquelas observações feitas continuamente no tempo, e discretas aquelas que são feitas em um intervalo de tempo específico.

Os modelos utilizados para descrever séries temporais são processos estocásticos, ou seja, processos controlados por leis probabilísticas.

O processo estocástico é definido como sendo de uma família $Z = \{Z(t), t \in T\}$, tal que, para cada $t \in T$, $Z(t)$ é uma variável aleatória. Nestas condições um processo estocástico é uma família de variáveis aleatórias que supomos definidas num mesmo espaço de probabilidade (Ω, A, P) . O conjunto T é tomado como o conjunto dos inteiros $Z = \{0, \pm 1, \pm 2, \dots\}$ ou seja os conjuntos dos reais \mathbb{R} . E para cada $t \in T$, $Z(t)$ ser uma variável aleatória real.

Sendo $t \in T, Z(t)$ uma variável aleatória definida sobre Ω , $Z(t)$ se torna uma função de dois argumentos, $Z(t, \omega), t \in T, \omega \in \Omega$.

A figura 2 mostra que para cada $t \in T$, temos uma variável aleatória $Z(t, \omega)$, com uma distribuição de probabilidade; é possível que a função densidade de probabilidade no instante t_1 seja diferente da função densidade de probabilidade no instante t_2 , para dois instantes t_1 e t_2 quaisquer, a situação usual é aquela em que a função densidade de probabilidade de $Z(t, \omega)$ é a mesma para todo $t \in T$.

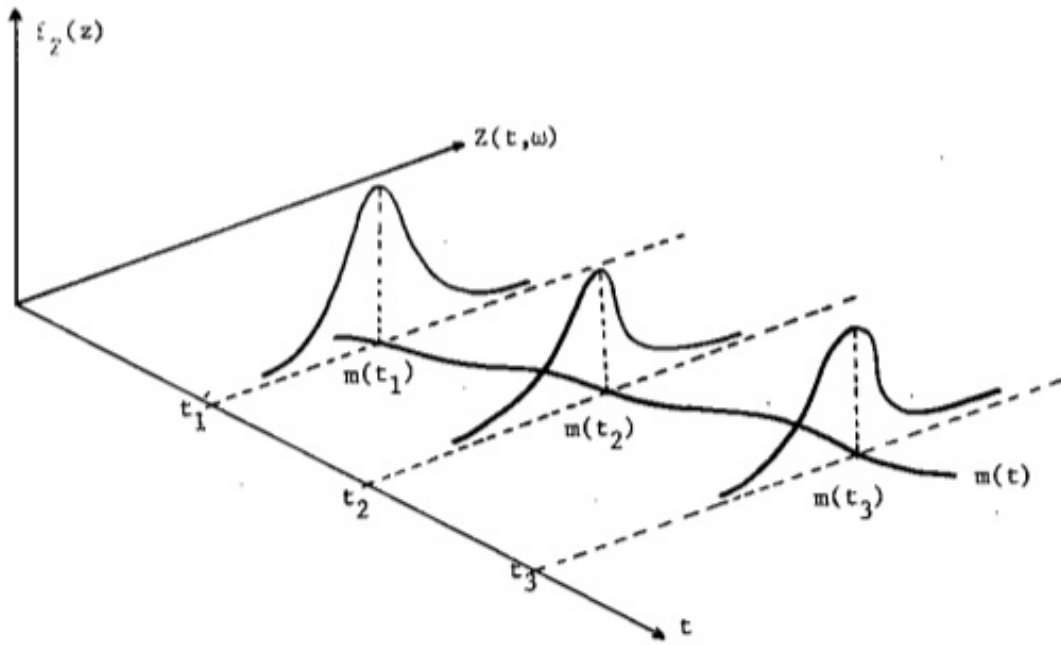


Figura 2. Um processo estocástico interpretado como uma família de variáveis aleatórias (MORETTIN e TOLOI, 2014, p. 20).

Por outro lado, para cada $\omega \in \Omega$ fixado, obteremos uma função de t , sendo uma realização ou trajetória do processo, ou ainda uma série temporal, como podemos ver na figura 3.

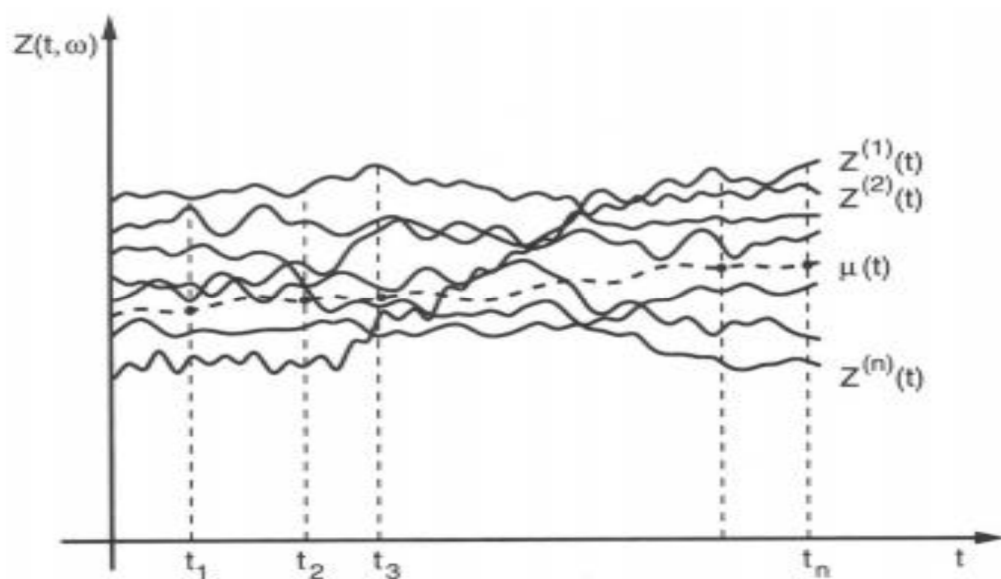


Figura 3. Um processo estocástico interpretado como uma família de trajetórias (MORETTIN e TOLOI, 2014, p. 20).

Vemos que as realizações de $Z(t, \omega)$ é dada por $Z^{(1)}(t), Z^{(2)}(t), \dots, Z^{(n)}(t)$, sendo o conjunto de todas as trajetórias ao longo do tempo, sendo chamado de *ensemble*. Cada trajetória particular de um processo estocástico é também chamada de série temporal ou função amostral. (SOUZA, 2015)

A análise de séries temporais tem como objetivo:

- (a) Investigar o mecanismo gerador da série temporal;
- (b) Fazer previsões de valores futuros da série, estas previsões podem ser a curto ou a longo prazo;
- (c) Descrever o comportamento da série, com o auxílio dos gráficos verificar a existência de tendência, ciclos e variações sazonais;
- (d) Procurar periodicidades relevante nos dados.

Em todos os casos, modelos probabilísticos ou modelos estocásticos são construídos, no domínio temporal ou de frequências, estes modelos devem ser simples e parcimoniosos (no sentido que os números de parâmetros envolvidos seja o menor possível (OLIVEIRA. 2012).

Podemos classificar os modelos para séries temporais em duas classes, segundo o número de parâmetros envolvidos:

- (a) Modelos paramétricos, para os quais o número de parâmetros é finito;
- (b) Modelo não-paramétricos, que envolvem um número infinito de parâmetros.

Nas classes de modelo paramétrico, a análise é feita no domínio do tempo. Dentre estes modelos os mais frequentemente usados são os modelos de erro (ou de regressão), os modelos autorregressivos e de médias móveis (ARMA), os modelos autorregressivos integrados e de médias móveis (ARIMA), modelos de memória longa (ARFIMA), modelos autorregressivos integrados de médias móveis sazonais (SARIMA).

O modelo autorregressivo de ordem p ($AR_{(p)}$) representado por;

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + \dots + \phi_p \tilde{Z}_{t-p} + a_t \quad (1.1)$$

em que, $Z_t = Z_{t-\mu}$ e μ é a média da série; ϕ_i são parâmetros do modelo a serem estimados com $i = \{1, 2, 3, \dots, p\}$; a_t representa o ruído que não pode ser explicado pelo modelo e supõe ser um ruído branco $a_t \sim N(0, \sigma_{a2})$.

Sendo definido o operador autorregressivo estacionário de ordem p como

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (1.2)$$

então pode-se escrever

$$\phi(B)\tilde{Z}_t = a_t \quad (1.3)$$

ou

$$\tilde{Z}_t = \phi \tilde{Z}_{t-1} + a_t \quad (1.4)$$

O processo de médias móveis de ordem q ($MA_{(q)}$) representado pela notação;

$$Z_t = \mu + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (2.1)$$

sendo $\tilde{Z}_t = Z_t - \mu$ e μ é a média da série: θ_i são parâmetros do modelo a serem estimados com $i = \{1, 2, 3, \dots, q\}$; a_t é ruído branco $a_t \sim N(0, \sigma_{a2})$.

$$\tilde{Z}_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t = \theta(B) a_t \quad (2.2)$$

onde

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (2.3)$$

é o operador de médias móveis de ordem q . Pode ser representado em forma reduzida por;

$$\tilde{Z}_t = a_t - \theta a_{t-1} \quad (2.4)$$

ou

$$\tilde{Z}_t = (1 - \theta B) a_t \quad (2.5)$$

Os modelos autorregressivos de média móveis são descritos como ARMA(p, q), AR(p) autorregressivo de ordem p , e MA(q) médias móveis de ordem q , podendo ser descrito da seguinte forma,

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \dots + \phi_p \tilde{Z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q} \quad (3.1)$$

e na forma reduzida de

$$\phi(B) \tilde{Z}_t = \theta(B) a_t \quad (3.2)$$

em que $\phi(B)$ e $\theta(B)$ são os operadores autorregressivos e de média móveis respectivamente.

O modelo autorregressivo integrado de médias móveis ARIMA(p, d, p), é um dos modelos mais utiliza na metodologia Box e Jenkins, pois ele descreve de maneira satisfatória series estacionárias e séries não-estacionárias, desde que elas não apresentem características explosivas. (SOUZA, 2015)

O modelo ARIMA(p, d, p), é definido através do operador autorregressivo p , operador diferenças d e operador médias móveis p , expressado a seguir por;

$$\phi(B)(1 - B)^d Z_t = \theta(B) a_t \quad (4.1)$$

ou

$$\phi(B) \Delta^d Z_t = \theta(B) a_t \quad (4.2)$$

em que d é o número de diferenças necessárias para tornar a série estacionária.

O modelo autorregressivo integrado de médias móveis sazonais $SARIMA(p, d, q)(P, D, Q)_S$, possui duas partes distintas: a primeira parte não-sazonal com parâmetros (p, d, q) , e a segunda parte sazonal com parâmetros $(P, D, Q)_S$, temos que p e P são os parâmetros autorregressivos das partes não-sazonal e sazonal respectivamente; d e D são o número de diferenças para as parte não-sazonal e sazonal; q e Q são os parâmetros médias móveis não-sazonal e sazonal; e S o período sazonal. (SOUZA, 2015)

O modelo $SARIMA(p, d, q)(P, D, Q)_S$, pode ser escrito como

$$\phi_p(B)\phi_P(B^S)(1-B)^d(1-B^S)^D Z_t = \theta_q(B)\theta_Q(B^S)a_t \quad (5.1)$$

ou de forma análoga,

$$\phi_p(B)\phi_P(B^S)\Delta^d\Delta_S^D Z_t = \theta_q(B)\theta_Q(B^S)a_t \quad (5.2)$$

no qual

$$\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \quad (5.3)$$

sendo o polinômio autorregressivo não-sazonal de ordem p ;

$$\phi_P(B^S) = (1 - \phi_1 B^S - \phi_2 B^{2S} - \dots - \phi_P B^{PS}) \quad (5.4)$$

sendo o polinômio autorregressivo não-sazonal de ordem P , estacionário;

$$\Delta^d = (1 - B)^d \quad (5.5)$$

sendo o operado de diferença não-sazonal e d indica o número de diferenças necessárias para retirar a tendência da série;

$$\Delta_S^D = (1 - B^S)^D \quad (5.6)$$

sendo o operado de diferença sazonal e D indica o número de diferenças sazonais necessárias para retirar a sazonalidade da série;

$$\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \quad (5.7)$$

sendo o polinômio médias móveis não-sazonal de ordem q ;

$$\theta_Q(B^S) = S(1 - \theta_1 B^S - \theta_2 B^{2S} - \dots - \theta_Q B^{QS}) \quad (5.8)$$

sendo o polinômio de médias móveis sazonal de ordem Q ;

Para obter um modelo, é preciso chegar aos parâmetros que compõem essa estrutura, Box e Jenkins possui uma metodologia chamada de ciclo iterativo, em que se os passos forem seguidos corretamente será obtido um modelo eficiente. A figura 2 traz a ilustração do ciclo iterativo.

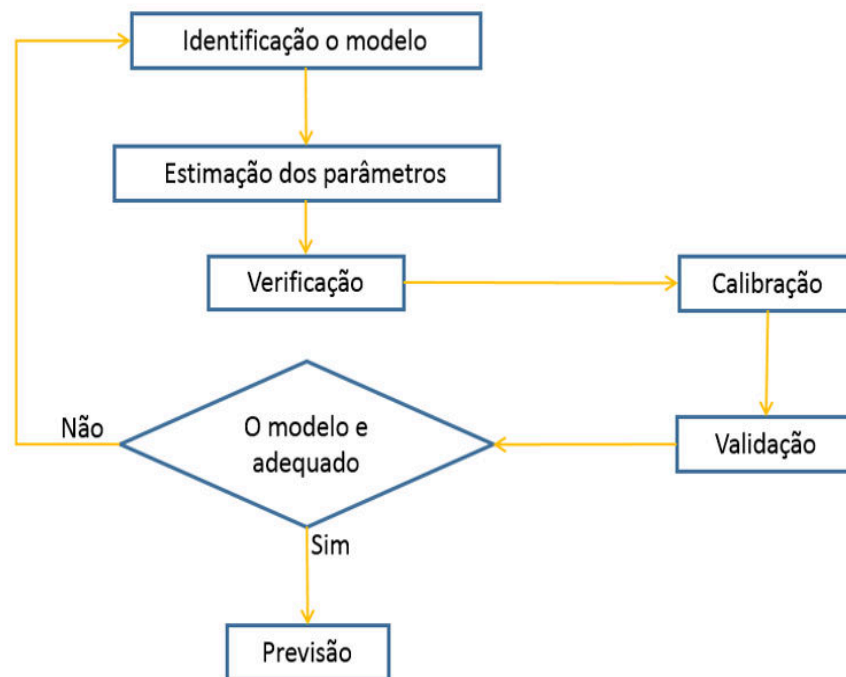


Figura 4. Ciclo iterativo de Box e Jenkins.

Fonte: Adaptado de Gujarati & Porter (2011, p. 772).

É proposto por Morettin e Toloi (2014) que o procedimento de identificação consiste em três partes:

- (a) Verificar se a série necessita de transformação nos dados originais, com objetivo de estabilizar a variância;

- (b) Aplicar a diferenciação na série até que a mesma se torne estacionária. O número de diferenças, d , é necessário para que os dados se tornem estacionário;
- (c) Estimar os valores de p e q , através da análise das autocorrelações (ACF) e autocorrelações parciais (PACF), serão apresentados os valores P e Q se a série apresentar sazonalidade, o número de termos média móvel e média móvel sazonal deve ser especificado através das ACF e PACF.

O primeiro passo na análise de séries temporais é identificar o modelo. A construção do gráfico da série original é necessária para a identificação algumas características, como tendência, sazonalidade, variabilidade, *outliers* (observações atípicas ou valores discrepantes).

Visualmente no gráfico é possível notar se existe a presença de tendência nos dados, para confirmar a existência de tendência e utilizado o teste Cox-Stuart (teste do sinal). O teste consiste em agruparmos as observações em pares $(Z_1, Z_{1+c}), (Z_2, Z_{2+c}), \dots, (Z_{N-c}, Z_N)$,

Onde;

$$c = \frac{N}{2} \text{ se } N \text{ for par}$$

$$c = \frac{N+1}{2} \text{ se } N \text{ for impar}$$

A cada par (Z_i, Z_{i+c}) associado o sinal $+$ se $Z_i < Z_{i+c}$ e o sinal $-$ se $Z_i > Z_{i+c}$, eliminando os empates.

Seja n o número de pares onde $Z_i \neq Z_{i+c}$, queremos testar

- $H_0: P(Z_i < Z_{i+c}) = P(Z_i > Z_{i+c}), \forall i$: não existe tendência;
- $H_1: P(Z_i < Z_{i+c}) \neq P(Z_i > Z_{i+c}), \forall i$: existe tendência;

Seja T_2 o número de pares com o sinal $+$, a hipótese H_0 é rejeitada caso $T_2 \geq n - t$, onde t é dado por;

$$t = \frac{1}{2} \left(n + \omega_{\alpha/2} \sqrt{n} \right) \quad (6.1)$$

sendo n o número de vezes que não houve empate na comparação entre Z_i e Z_{i+c} e $\omega_{\alpha/2}$ é o valor tabelado ao nível significância α . Caso $n \leq 20$, o valor tabelado de $\omega_{\alpha/2}$ é dado da

distribuição binomial com $p = 0,5$ e, caso $n > 20$, utiliza-se a aproximação normal. (SOUZA, 2015)

Para identificar a sazonalidade e utilizado o teste Fisher, esse teste é utilizado para confirmar a existência da sazonalidade na série temporal Z_t , o teste Fisher testa as seguintes hipóteses:

- H_0 : não existe sazonalidade nos dados.
- H_1 : existe sazonalidade nos dados.

Supondo que o processo Z_t seja gaussiano, baseado na estatística;

$$g = \frac{\max I_j^{(N)}}{\sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} I_j^{(N)}} = \frac{I^1}{\sum_{j=1}^{\lfloor \frac{N}{2} \rfloor} I_j^{(N)}} \quad (7.1)$$

Para N ímpar, a distribuição exata de g , sobe H_0 , é dada por:

$$P(g > a) = n (1 - a)^{n-1} - \binom{n}{2} (1 - 2a)^{n-1} + \dots + (-1)^x \binom{n}{x} (1 - xa)^{n-1} \quad (7.2)$$

em que $n = \left\lfloor \frac{n}{2} \right\rfloor$ e x é o maior inteiro menor que $\frac{1}{a} \left(x = \left\lfloor \frac{n}{2} \right\rfloor \right)$.

Assim, para um dado nível de significância α , podemos utilizar a equação (7.2) para encontrar o valor crítico α tal que $P(g > a(\alpha)) = \alpha$. Se o valor observado da estatística (g_{obs}) for maior que α , rejeita-se H_0 para o nível de significância estabelecido, indicando que há sazonalidade na série temporal.

Após a identificação dos modelos é preciso estimar os parâmetros, O objetivo de estimar os parâmetros é determinar os valores de p , d , q do modelo ARIMA (p , d , q) e também (P, D, Q)s, quando a série apresentar a componente sazonal (SARIMA).

Com o modelo e os parâmetros escolhidos, e feito a verificação para a escolha do melhor modelo, como critério de decisão é usado o critério de informação de Akaike (AIC) e erro absoluto médio percentual (MAPE).

O critério de AIC tem como base comparar modelos construídos na maximização do logaritmo da função de verossimilhança, penalizando os modelos com os maiores parâmetros. Dessa forma quanto menor for o valor do AIC de um modelo melhor será o modelo. (SOUZA, 2015)

O critério de informação de Akaike é representado por;

$$AIC = -2\ln(L) + 2m \quad (8.1)$$

definindo L como sendo a máxima verossimilhança do modelo ajustado é por m o número de parâmetros ajustados.

O erro absoluto médio percentual é representado por;

$$MAPE = \frac{1}{h} \sum_1^h \left| \frac{e_t(h)}{Z_{t=h}} \right| \times 100 \quad (9.1)$$

Com o objetivo de ter a previsão o mais próximo dos valores observados, esses indicadores medem o desempenho dos modelos em suas previsões, assim quanto menor o valor do erro de previsão melhor é a eficácia do modelo.

Desta forma esses critérios são usados para determinar os melhores modelos, assumindo que os modelos com os menores valores de AIC e MAPE são considerados os melhores modelos para previsão.

A normalidade é indispensável para a validação dos modelos, são usados alguns testes para averiguação da normalidade, entre eles temos o teste Shapiro Wilk com grau de significância (α) igual a 0,05 o teste testa as seguintes hipóteses;

- H_0 : A variável aleatória adere à distribuição Normal;
- H_1 : A variável aleatória não adere à distribuição Normal

Se o p-valor do teste for menor que o nível de significância escolhido, rejeita-se a hipótese de normalidade.

O Teste de Kolmogorov-Smirnov é um teste de aderência. Verifica o grau de concordância entre distribuições de um conjunto de valores amostrais e alguma distribuição teórica, assim verifica se os dados seguem uma distribuição normal. (SOUZA, 2015)

As hipóteses do teste são as seguintes;

- H_0 : A amostra provém de uma distribuição teórica específica (distribuição normal);
- H_1 : A amostra não provém de uma distribuição teórica específica (distribuição normal).

O teste observa a máxima diferença entre a função de distribuição acumulada assumida para os dados, como critério, compara-se esta diferença com um valor crítico, para um dado nível de significância. Se o resultado do teste for maior que o valor crítico do nível de significância não se rejeita a hipótese de normalidade.

Na aplicação dos testes de Box-Pierce e teste de Ljung-Box o teste de hipóteses que é conjunto, é representado pelas hipóteses a seguir:

- $H_0: \hat{r}_1 = \hat{r}_2 = \dots = \hat{r}_k = 0$
- $H_1: \hat{r}_1 = \hat{r}_2 = \dots = \hat{r}_k \neq 0$

Simplificando quando o p-valor dos testes Box-Pierce e Teste de Ljung-Box é maior que o nível de significância usado, não se rejeita a hipóteses de normalidade, assim tendo um modelo adequado.

Todas as análises estatísticas e todos os gráficos serão feitos no software R. 3.4.1 (R DEVELOPMENT CORE TEAM, 2015) o software R é um software livre todas a análises estão disponíveis em procedimentos disponíveis em pacotes dedicado análise de séries temporais. Para as análises dos testes estatísticos, será considerado o nível de significância usual de 0,05 para todos os testes.

3 RESULTADOS E DISCUSSÕES

As figuras 5 e 6 apresentam os gráficos das séries de dados originais referente à precipitação pluviométrica da cidade de Porto Velho (PVH) e precipitação pluviométrica da cidade de Vilhena (VHA).

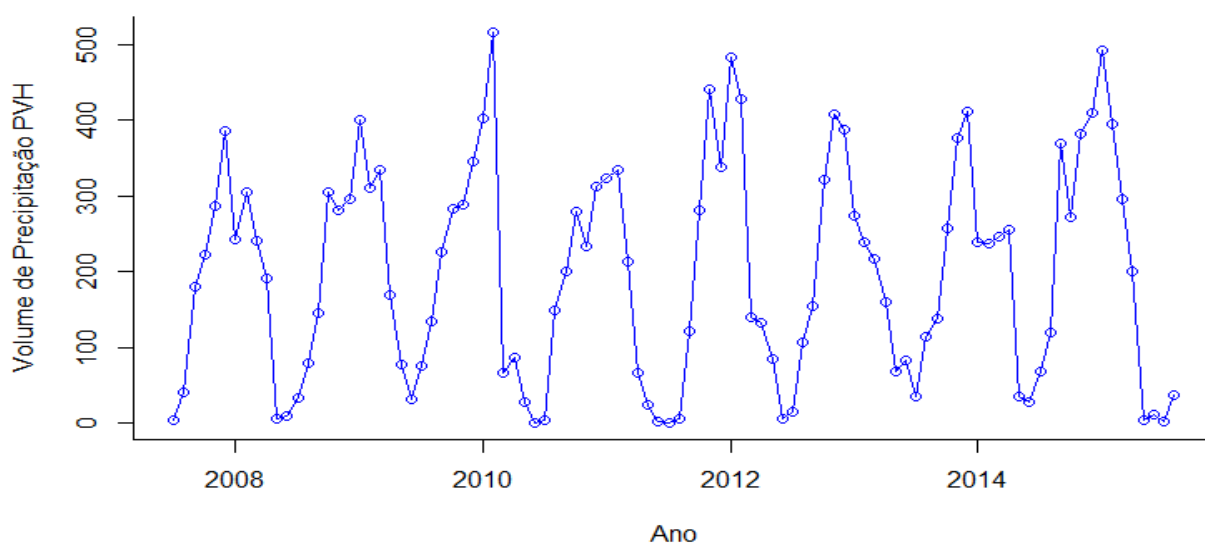


Figura 5. Representação gráfica da série original referente a precipitação pluviométrica da cidade de Porto Velho RO, no período de Julho de 2007 até Agosto de 2015.

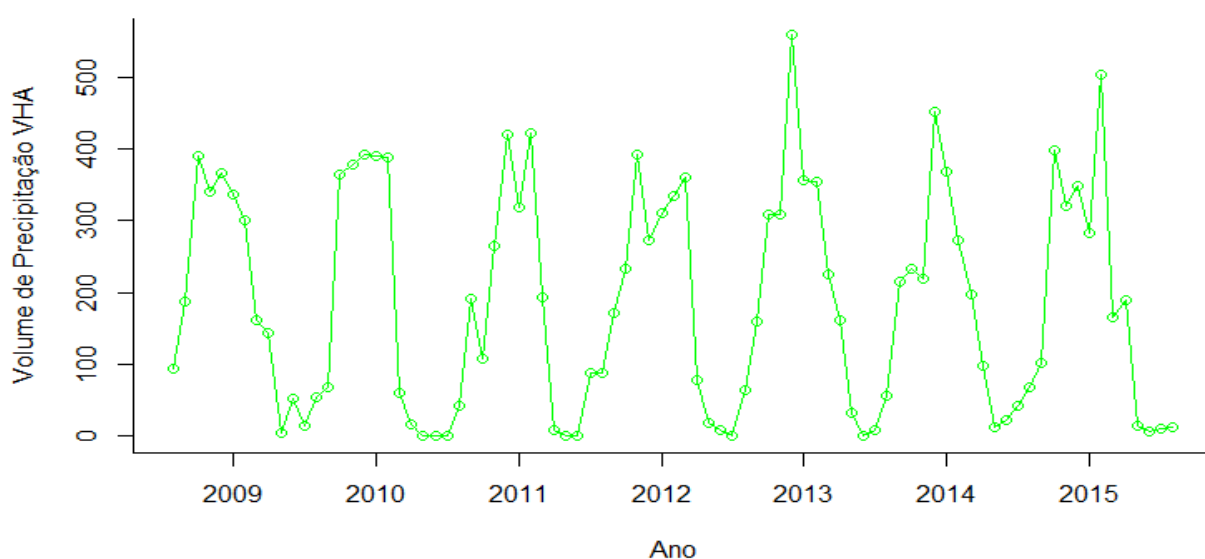


Figura 6. Representação gráfica da série original referente a precipitação pluviométrica da cidade de Vilhena RO, no período de Agosto de 2008 até Agosto de 2015.

Nota-se nas figuras 5 e 6 que os dados possuem um comportamento irregular com altos e baixos no volume da precipitação no ano. Os dados têm o comportamento similar nas duas cidades, com um grande índice de precipitação nos meses de Dezembro a Fevereiro e um volume muito pequeno nos meses de Junho a Agosto, chegando a ter meses com volume de precipitação próximos a zero. Esse comportamento indica presença de sazonalidade nos dados com período de 12 meses.

As figuras 7 e 8 apresentam os gráficos de BoxPlot para identificação de *outliers*, os *outliers* são valores discrepantes ou valores atípicos que tem um grande afastamento dos demais valores da série, em alguns casos os *outliers* tem influência negativa na análise dos dados pois eles podem afetar os resultados.

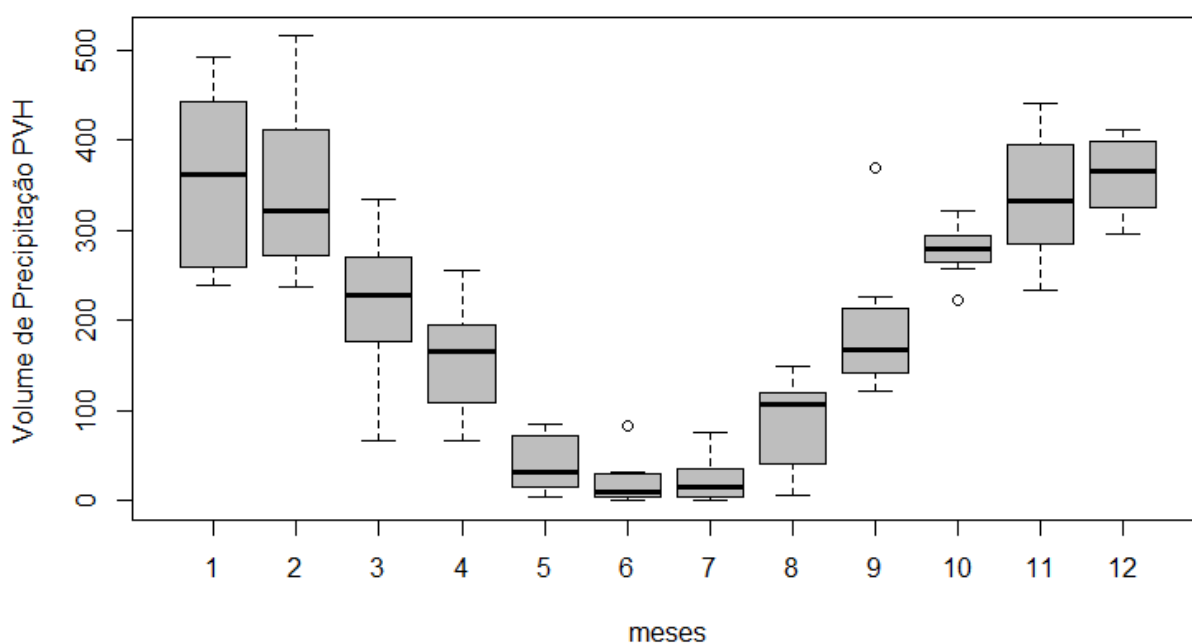


Figura 7. Gráfico BoxPlot apresentando os dados em médias mensais das séries originais da precipitação pluviométrica da cidade de Porto Velho RO.

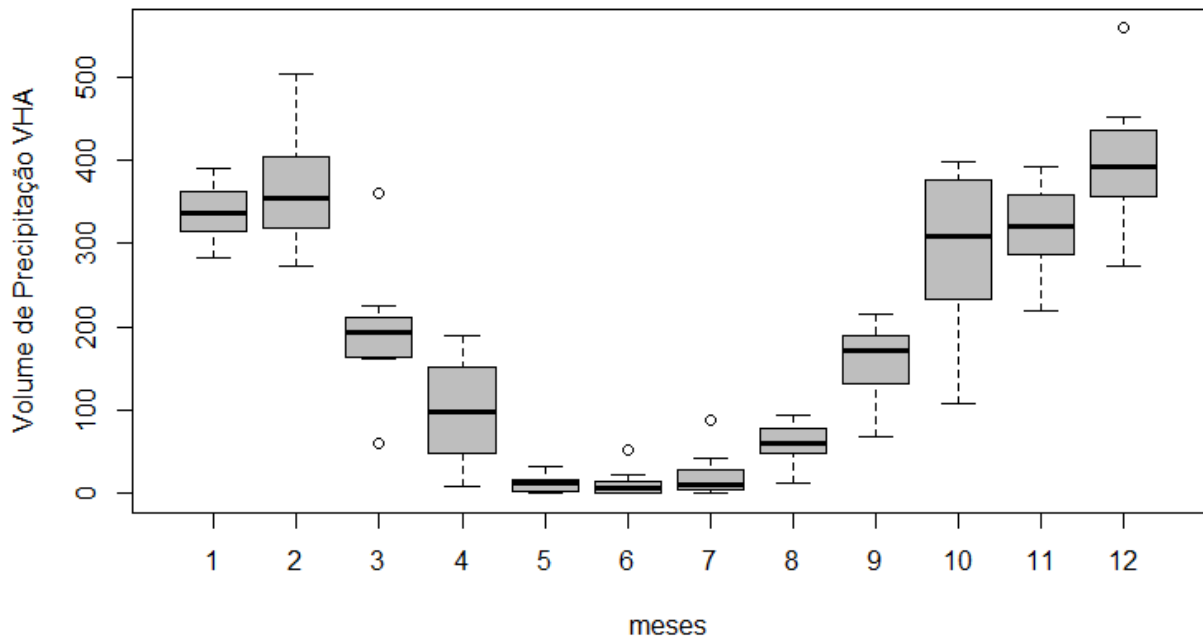


Figura 8. Gráfico BoxPlot apresentando os dados em médias mensais das séries originais da precipitação pluviométrica da cidade de Vilhena RO.

Nota-se que nas figuras 7 e 8, existe *outliers* em alguns meses dos dados das duas cidades em estudo, esses *outliers* representam meses que tiveram o volume de precipitação anormal quando comparado com os mesmos meses dos outros anos. O número de *outliers* nos dados podem ser considerados pequenos, pois não são todos os meses que possuem *outliers*, e nos meses que tem *outliers* só possui um em cada, assim não exercendo grande influência nas análises dos dados.

O gráfico de BoxPlot nos dá uma confirmação que o período sazonal é de 12 meses. Pode-se notar que nas cidades de Porto Velho e Vilhena a precipitação pluviométrica segue o mesmo comportamento tendo os seus menores índices nos meses de Maio a Julho. É possível notar que a cidade de Porto Velho possui um volume de precipitação um pouco maior que a cidade de Vilhena, mas isso não afeta o comportamento da precipitação nas duas cidades.

As figuras 9 e 10, apresentam a decomposição das séries originais da precipitação pluviométrica das cidades de Porto Velho e Vilhena respectivamente, onde a série original é separada em quatro partes: a série original: a parte tendenciosa: a parte sazonal e por último o ruído da série.

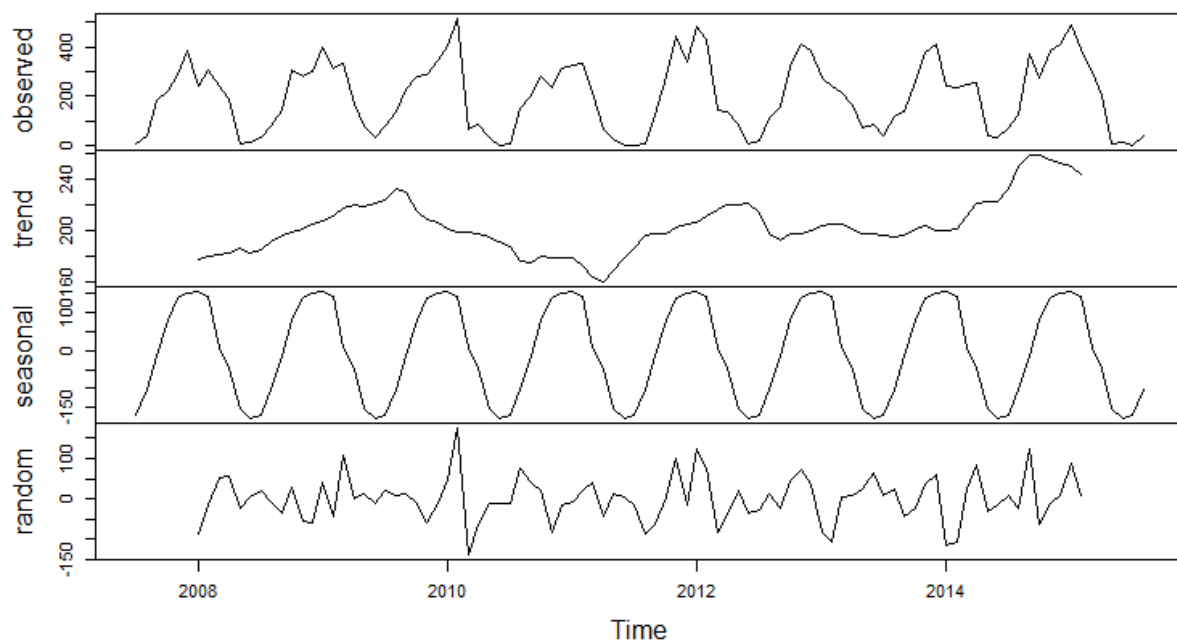


Figura 9. Gráfico da decomposição das series da precipitação pluviométrica da cidade de Porto Velho.

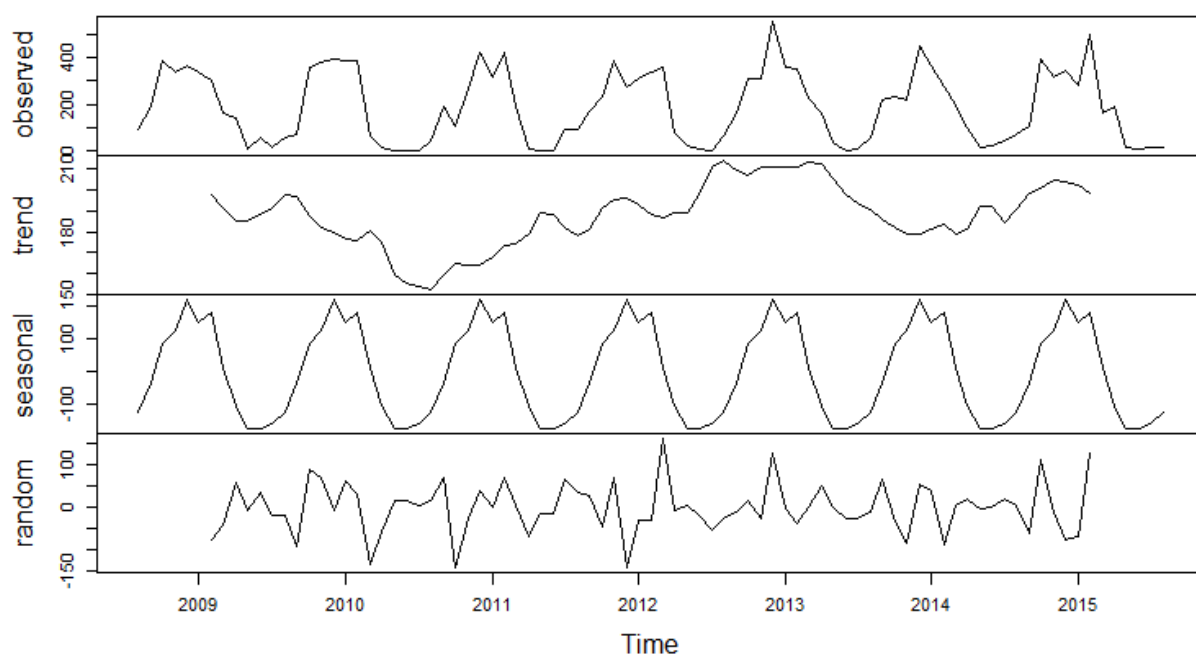


Figura 10. Gráfico da decomposição das series da precipitação pluviométrica da cidade de Vilhena.

Pode-se observar que na parte de tendência (trend) nos dois gráfico que não há indício que indique tendência nos dados, pois a linha da tendência não possui nenhum comportamento padronizado, já na parte sazonal (seasonal) dos gráficos pode-se notar que a sazonalidade é bem expressiva, pois o mesmo comportamento se repete a cada ano como é notado na linha da

sazonalidade que tem picos há cada ano, para ter certeza da existência da tendência e da sazonalidade ser realizado testes específicos para tendência e sazonalidade.

Para testa se existe a presença de sazonalidade e tendência nos dados será realizado o teste Fisher para identificar sazonalidade e o teste Cox-Stuart para verificar a existência de tendência.

Será realizado o teste Fisher para identificar se existe sazonalidade nos dados. Em que hipótese nula é a não existe sazonalidade para um nível significância de 0,05. Então se o p-valor for igual ou menor que 0,05 os dados possuem sazonalidade.

Para o teste Cox-Stuart a hipótese nula é a não existência de tendência nos dados, para que essa hipótese seja rejeitada ao nível significância de 0,05, o p-valor tem que ser igual ou menor que 0,05.

A tabela 1 traz os resultados dos testes Fisher e Cox-Stuart da precipitação pluviométrica das cidades de Porto Velho e Vilhena.

Tabela 1. Resultados dos testes de sazonalidade e tendência para a precipitação pluviométricas das cidades de Porto Velho e Vilhena.

Teste Fisher		Teste Cox-Stuart
Variável	p-valor	p-valor
Precipitação PVH	6,3820E-20	0,2200
Precipitação VHA	2,5960E-23	0,5612

A tabela 1, traz os resultados dos p-valores dos testes Fisher e Cox-Stuart, nota-se que para a precipitação pluviométrica das duas cidades o teste Fisher apresenta um p-valor menor que 0,05, sendo assim rejeita-se a hipótese nula. O teste Cox-Stuart apresenta p-valores maior que 0,05 para as duas cidades, sendo assim não é rejeitada a hipótese nula (não existe tendência nos dados). Após a confirmação da existência de sazonalidade nos dados, será necessário realizar a diferenciação na defasagem sazonal para deixar os dados sem sazonalidade.

Na figura 11, tem-se os correlogramas das funções de autocorrelação (ACF) e a autocorrelação parcial (PACF) das séries de Precipitação pluviométricas de Porto Velho e Vilhena.

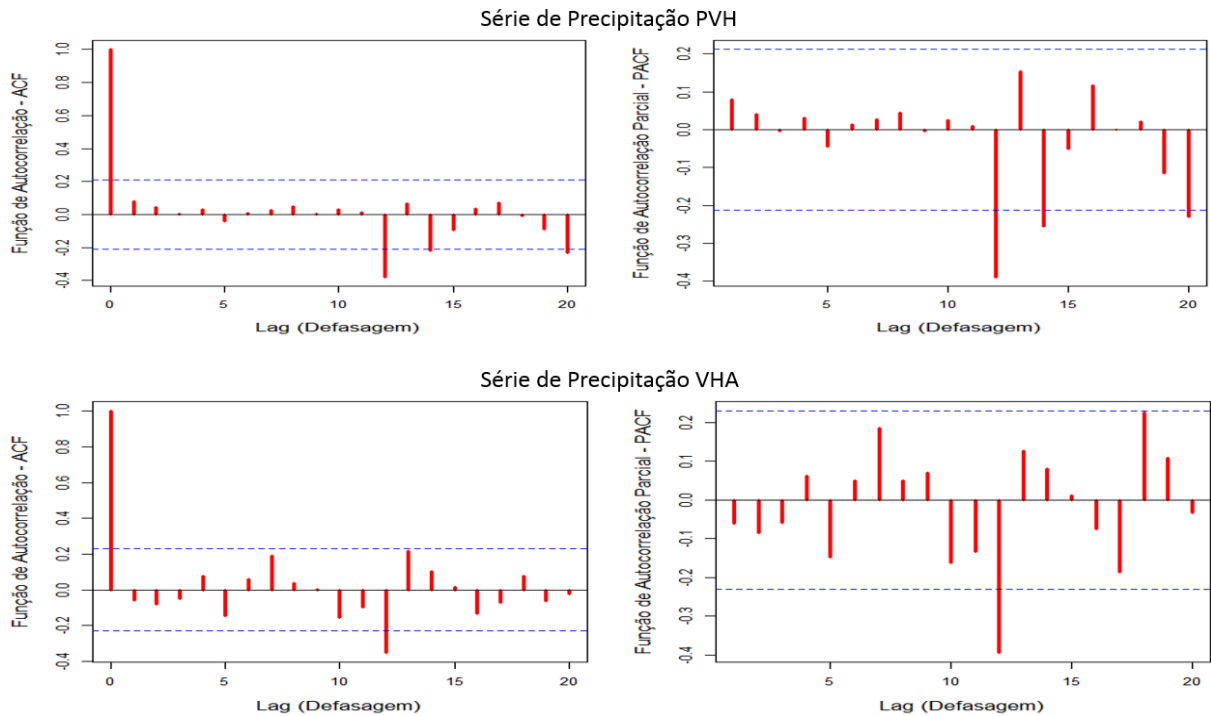


Figura 11. Representação gráfica das funções de autocorrelação (ACF) e autocorrelação parcial (PACF) da série diferenciada sazonalmente da precipitação pluviométrica das cidades de Porto Velho Vilhena.

Os correlogramas são usados como auxílio visual para facilitar a aplicação da metodologia Box e Jenkins facilitando assim a escolha dos parâmetros, ajustando o modelo SARIMA $(p, d, q) \times (P, D, Q)_S$.

Para a seleção dos melhores modelos foram usados os critérios de AIC e MAPE, os valores são apresentados nas tabelas 3 e 4, os dois modelos selecionados foram os que apresentaram os menores valores de cada critério. Os modelos selecionados para a Precipitação PVH e a Precipitação VHA foram:

- SARIMA $(1,0,2) \times (1,0,1)_{12}$ Precipitação PVH
- SARIMA $(1,0,1) \times (1,0,1)_{12}$ Precipitação VHA

A metodologia Box e Jenkins possui alguns requisitos para que os modelos sejam considerados adequados para previsões, entre eles está o teste dos resíduos, é necessário testar os resíduos para saber se eles são independentes e possuem normalidade, para testar a normalidade dos resíduos será usado os testes de Shapiro-Wilk, Kolmogorov-Smirnov e os

testes conjuntos Box-Pierce e Ljung-Box, os testes serão realizado considerando um nível de significância de $\alpha = 0,05$. O resultado dos testes está exposto na tabela 2.

Tabela 2. Resultados dos testes Shapiro-Wilk, Kolmogorov-Smirnoff, Box-Pierce e Ljung-Box para os resíduos dos modelos escolhidos.

Teste dos Resíduos					
Modelos SARIMA $(p, d, q)x(P, D, Q)_{12}$		p-valor			
		Shapiro-Wilk	Kolmogorov-Smirnoff	Box-Pierce	Ljung -Box
Precipitação PVH	$(1,0,2)x(1,0,1)_{12}$	0,0565	0,5894	0,5407	0,302
Precipitação VHA	$(1,0,1)x(1,0,1)_{12}$	0,0523	0,5000	0,3582	0,1256

Observa-se que todos os p-valores possuem um valor maior que 0,1, segundo os testes de hipóteses dos testes Shapiro-Wilk, Kolmogorov-Smirnoff, Box-Pierce e Ljung-Box, para todos os testes descritos a hipótese nula (H_0 : A variável aleatória adere à distribuição Normal) é igual. Sendo assim com p-valor maior que 0.05 não se rejeita a hipótese H_0 , com base nos testes feitos pode-se afirmar que os resíduos dos modelos proposto para a precipitação pluviométrica das cidades de Porto Velho e Vilhena seguem uma distribuição Normal.

Após testa a normalidade dos resíduos é necessário verificar se os parâmetros escolhidos para cada modelo são adequados, as tabelas 3 e 4 apresentam as estimativas dos parâmetros dos modelos escolhidos da precipitação pluviométrica das cidades de Porto Velho e Vilhena.

Tabela 3. Estatísticas do modelo SARIMA (p,d,q) x (P,D,Q)s, da precipitação pluviométrica da cidade de Porto Velho.

SARIMA (1,0,2) x (1,0,1) ₁₂						
Parâmetros	Estimativa	Erro Padrão	Intervalo de confiança		Critérios	
			2,50%	97,50%	AIC	MAPE
AR1	0,9525	0,0459	0,86	1,04	1136,03	3,671
MA1	-0,7047	0,0977	-0,90	-0,51		
MA2	-0,2953	0,0976	-0,49	-0,10		
SAR1	0,9997	0,0013	0,99	1,00		
SMA1	-0,9628	0,0918	-1,14	-0,78		

Tabela 4. Estatísticas do modelo SARIMA (p,d,q) x (P,D,Q)s, da precipitação pluviométrica da cidade de Vilhena.

SARIMA (1,0,1) x (1,0,1) ₁₂						
Parâmetros	Estimativa	Erro Padrão	Intervalo de confiança		Critérios	
			2,50%	97,50%	AIC	MAPE
AR1	0,9305	0,0462	0,84	1,02	993,09	2,6135
MA1	-1,0000	0,0046	-1,00	-0,99		
SAR1	0,9999	0,0006	0,99	1,00		
SMA1	-0,9686	0,0927	-1,15	-0,78		

As tabelas 3 e 4 apresentam os resultados dos parâmetros de cada modelo escolhido, para verificar se os parâmetros são adequados, é usado o intervalo de confiança como teste. Sendo assim o parâmetro é considerado adequado se o 0 não estiver incluso no intervalo de confiança. Nota-se que todos os parâmetros escolhidos possuem o intervalo de confiança adequado, pois o 0 não está contido em nenhum deles.

Após as análises dos modelos e seus parâmetros, foi verificado que todos os requisitos da modelagem Box e Jenkins foram cumpridos, dessa forma será dada continuidade ao trabalho, em que o próximo passo é a previsão da série.

Na figura 12, é apresentado o gráfico da previsão da série de precipitação pluviométrica da cidade de Porto Velho RO.

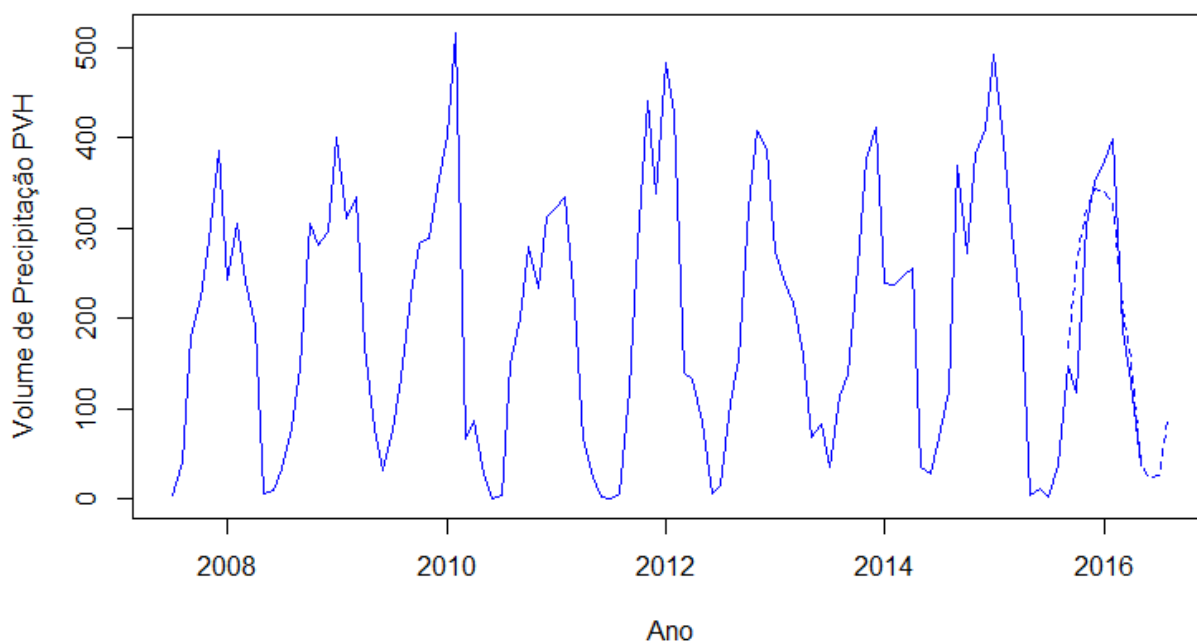


Figura 12. Gráfico da previsão da série de precipitação pluviométrica da cidade de Porto Velho RO, no período de Setembro de 2015 a Maio de 2016.

A figura 12, apresenta a previsão da precipitação pluviométrica da cidade de Porto Velho, a linha contínua representa os dados originais que são de Julho de 2007 a Maio de 2016, a previsão foi feita nos últimos nove meses da série original, esses meses foram deixado de fora de todas as análise anteriores e usado somente para a validação da previsão, que está expressa pela linha tracejada nos meses de Setembro de 2015 a Maio de 2016. Visualmente e possível notar que a previsão teve um comportamento similar ao da série original, nota-se que a um pequeno desvio na previsão, possivelmente com algum valore previsto estando a baixo dos valores reais, visualmente no gráfico da previsão é possível dizer que a modelo é considerada bom. Podemos averiguar melhor na tabela 5.

A figura 13 apresentado o gráfico da previsão da série de precipitação pluviométrica da cidade de Vilhena RO.

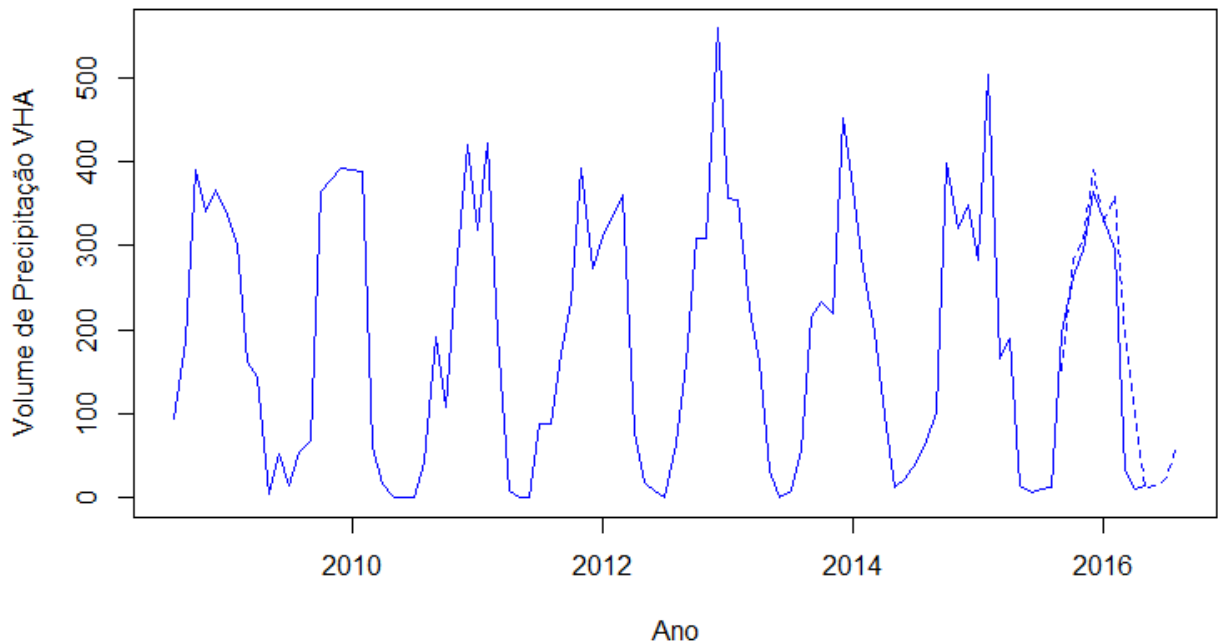


Figura 13. Gráfico da previsão da série de precipitação da cidade de Vilhena RO, no período de Setembro de 2015 a Maio de 2016.

Na figura 13, temos o gráfico dos dados originais do período de Agosto de 2008 a Maio de 2016, assim como não previsão do modelo anterior a linha tracejada representa a previsão da precipitação da cidade de Vilhena, no período de Setembro de 2015 Maio de 2016, nota-se que em quase todo período previsto a linha tracejada está em cima da linha continua que representa a série original, ocorreu apenas um pico na previsão, mas o comportamento da série prevista e da série original estão semelhantes, visualmente a previsão se mostro eficiente, os valores da previsão da cidade de Vilhena está exposto na tabela 6.

As tabelas 5 e 6 trazem os valores previstos e valores reais das previsões da precipitação pluviométrica das cidades de Porto Velho e Vilhena.

Tabela 5. Valores reais e valores previstos do modelo SARIMA(1,0,2) x (1,0,1)₁₂, no período de Setembro de 2015 a Maio de 2016, da precipitação pluviométrica da cidade de Porto Velho.

Período	Limite Inferior 5%	Previsão	Limite Superior 95%	Valor Original	Erro de previsão
set/15	101,87	147,00	233,95	167,91	-20,91
out/15	193,63	118,60	329,64	261,63	-143,03
nov/15	251,68	302,60	387,91	319,80	-17,20
dez/15	274,69	350,50	411,12	342,90	7,60
jan/16	271,56	374,45	408,17	339,87	34,58
fev/16	260,68	398,40	397,45	329,06	69,34
mar/16	141,52	187,20	278,43	209,98	-22,78
abr/16	83,53	128,60	220,57	152,05	-23,45
mai/16	-27,40	37,80	109,76	41,18	-3,38

A tabela 5 apresenta os valores da previsão da precipitação de Porto Velho, junto com os intervalos de confiança, valores originais e erro de previsão, pode se notar que com exceção de dois meses o restante está dentro do intervalo de confiança, observando a coluna seis da tabela 5 nota-se que existem dois valores que se destacam dos demais, esses valores são a diferença entre os valores originais e valores previstos ele são atribuído aos meses de Outubro e Fevereiro, nota-se que os dois meses que se destacarão são os únicos que possuem os valores fora do intervalos de confiança.

O mês de Outubro teve um valor previsto bem a baixo do limite inferior do intervalo de confiança, justificando o comportamento que foi visto anteriormente na figura 12, em que a linha da previsão tem um pequeno desvio, o mês de Fevereiro está com apenas um unidade de medidas a cima do limite do intervalo, se desconsiderarmos esse desvio do mês de Outubro é possível observar que prevista tem o mesmo comportamento da série dos valores original.

Tabela 6. Valores reais e valores previstos do modelo SARIMA(1,0,1) x (1,0,1)_12, no período de Setembro de 2015 a Maio de 2016, da precipitação pluviométrica da cidade de Vilhena.

Período	Limite Inferior 5%	Previsão	Limite Superior 95%	Valor Original	Erro de previsão
set/15	85,59	195,15	216,86	151,22	43,93
out/15	216,37	263,64	347,96	282,17	-18,53
nov/15	242,58	295,40	374,44	308,51	-13,11
dez/15	324,57	363,80	456,66	390,61	-26,81
jan/16	262,74	331,80	395,03	328,89	2,91
fev/16	292,99	294,60	425,45	359,22	-64,62
mar/16	124,28	32,60	256,89	190,58	-157,98
abr/16	31,86	10,40	164,59	98,23	-87,83
mai/16	53,40	15,40	79,44	13,02	2,38

Analisando a tabela 6 nota-se que os valores previstos da precipitação da cidade de Vilhena foi próximo dos valores reais. Os seis primeiros meses tiveram todos os valores dentro do intervalo de confiança e com erros pequenos, os meses de Março e Abril tiveram seus valores previsto a baixo do limite inferior dos intervalos de confiança, é comum para algumas análises de previsão que após os primeiros valores previstos se distanciem dos valores reais, essa são análises para curto prazo, no geral foi considerado que a previsão da precipitação da cidade de Vilhena foi eficiente.

4 CONCLUSÕES

O presente trabalho teve como objetivo estudar o comportamento da Precipitação Pluviométrica nas cidades de Porto Velho e Vilhena, e testar a metodologia Box e Jenkins, para a previsão. As cidades em estudo ficam a 707 quilômetros de distância uma da outra, sendo cada uma em uma ponta do estado de Rondônia.

Observou que a precipitação pluviométrica tem um comportamento similar nas duas cidades, foi observado que nos meses de Outubro a Fevereiro o volume da precipitação é maior que nos outros meses, isso se dá pelo período de chuva no estado de Rondônia, que dura aproximadamente seis meses.

A metodologia Box e Jenkins se mostrou eficiente para previsão da precipitação das cidades de Porto Velho e Vilhena. Concluiu-se que a metodologia pode ser usada nos dois extremos do estado de Rondônia, sendo eficiente descrever o comportamento e fazer previsões da precipitação pluviométrica.

5 REFERÊNCIAS BIBLIOGRÁFICAS

BOX, G. E. P., JENKINS, G. M., REINSEL, G. C., (1994) **Time Séries Analysis Forecasting and Control**, Prentice Hall, New Jersey, 3rd edition.

CARVALHO R. L. S.; NASCIMENTO B. I. S.; QUERINO C. A. S.; SILVA M. J. G.; DELGADO A. R. S. **Comportamento das series temporais de temperatura do ar, umidade e precipitação pluviométrica no município de Ariquemes (Rondônia-Brasil)**. Revista Brasileira de Climatologia. Vol. 18, p 123-142 JAN/JUN 2016.

GABRIEL CONSTANTINO BLAIN. **Considerações estatísticas relativas a oito series de precipitação pluvial da secretaria de agricultura e abastecimento do estado de São Paulo**. Revista Brasileira de Meteorologia, v.24, n.1, 12-23, 2009.

GUJARATI, D. N.; PORTER, D. C. **Econometria básica**. 5 ed. Porto Alegre: AMGH, 2011. 59

INSTITUTO NACIONAL DE METEOROLOGIA.NOTA TECNICA NÚMERO 001/2011/SEGER/LAIME/CSC/INMET: **Rede de Estações Meteorológicas Automáticas do INMET**, 2011. 11 p.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. São Paulo: Edgard Blucher, 2014.

OLIVEIRA, J. S. **Desenvolvimento e treinamento de redes neurais artificiais para processamento de dados de radiação solar**. Trabalho de Conclusão de Curso (Graduação em Engenharia de Energia) - Universidade Federal de Santa Catarina, Araranguá, 2017.

OLIVEIRA, P. J. A. **Previsão de demanda de água em sistemas de abastecimento através de modelos arima sazonais**. 2012. Dissertação de Mestrado.

PEREIRA A. R.; COSTA A. S.; OLIVEIRA V. G.; BORGES P. F.; FILHO A. I. **Análise do comportamento das médias anuais da precipitação pluvial e temperatura da cidade de Areia Paraíba**. GAIA SCIENTIA. Vol 9 p 67-73 2015.

PINTO W. P.; LIMA G. B.; ZANETTI J. B. **Estudo de predição e previsão de temperaturas medias diária da cidade de Cariacica, Espírito Santo, Brasil, utilizando a metodologia de series temporais**. GEINTEC Vol. 5/n. 1/ p.1881-1895 2015.

R DEVELOPMENT CORE TEAM (2015). **R: a language and environment for statistical computing**. R Foundation for Statistical Computing,Vienna, Austria. Disponível em: < <http://www.r-project.org> >. Acesso em: out de 2015.

SILVA M. I, GUIMARÃES E. C, TAVARES, M. 2008. **Previsão da temperatura média mensal de Uberlândia, MG, com modelos de séries temporais.** Revista Brasileira de Engenharia Agrícola e Ambiental, 12(5):480–485.

SOUZA R. B. **Análise De Séries Temporais E Previsão: Um Estudo De Aplicação Prática Da Metodologia De Box & Jenkins No Consumo Mensal De Energia Elétrica Da Região Norte Do Brasil.** Trabalho de Conclusão de Curso (Graduação em Estatística) – Universidade Federal de Rondônia.

TABARI, H.; TALAEI, H. P. **Sensitivity of evapotranspiration to climatic change in different climates.** Global and Planetary Change, v.115, p. 16–23, 2014.

VON RANDOW, R.C.S.; VON RANDOW, C.; HUTJES, R. W. A.; TOMASELLA, J.; KRUIJT, B. **Evapotranspiration of deforested areas in central and southwestern Amazonia.** Theoretical and Applied Climatology. v. 109, n. 1-2, p. 205-220. 2012.

6 APÊNDICE A - ROTINA DA ANÁLISE DE SÉRIES TEMPORAIS

```
##script precipitação PVH
library(tseries)#ativa o pacote
library(TSA)#ativa o pacote
getwd()
setwd("C:/Users/Cliente/Downloads/Nova pasta/dados e script TCC")
getwd()
x <- read.table("txt.prc.pvh.txt", header=T)
x
x=ts(x,frequency=12,start=c(2007,7))
x1 = window(x, start=c(2007, 7), end=c(2015, 8))
class(x); start(x); end(x); frequency(x)
class(x1); start(x1); end(x1); frequency(x1)
win.graph(width=8,height=5,points=12)
plot(x1,xlab='Ano',ylab='Precipitação PVH',main="",col='blue', type='o', bty="l")
win.graph(width=8,height=5,points=12)
plot(decompose(x1))
## Teste Cox-Stuart - verificar a tendência da série
cox.stuart=function(x, trend=c("+","-")){
  trend=match.arg(trend)
  N=length(x)
  p1=x[1:(N/2)]
  p2=x[(N/2+1):N]
  sinais=sign(p2-p1)
  if(trend=="-") T2=sum(sinais<0) else T2=sum(sinais>0)
  n=length(sinais[sinais!=0])
  if(trend=="+") tcrit=qbinom(.95,n,.5) else tcrit=qbinom(.05,n, .5)
  if(trend=="+") pv=1-pbinom(T2, n,.5) else pv=pbinom(T2, n, .5)
  if(round(pv, 4)==0) pv="<0.0001" else pv=round(pv, 4)
  data.frame("N"=N, "n"=n, "estatistica"=T2, "valor.critico"=tcrit, "tendencia"=trend, "p-
valor"=pv)}
cox.stuart(x1,trend="+")
##Teste Fisher - verificar a sazonalidade da série
win.graph(width=8,height=5,points=10)
pgram=spectrum(x1, fast=FALSE, detrend=FALSE, log="no", plot=T, main="Periodograma
da série")$spec
n=length(pgram)
pvalue.g=function(g,obs, n) n*(1-g.obs)^(n-1)
g.obs=max(pgram)/sum(pgram)
pvalue.g(g.obs,n)
win.graph(width=8,height=5,points=12)
layout(1:2)
m = acf(x1, plot=F); m$lag = m$lag*12; plot(m, xlab='Defasagem',ylab='Autocorrelação',
main="")
m1 = pacf(x1, plot=F); m1$lag = m1$lag*12; plot(m1,
xlab='Defasagem',ylab='Autocorrelação Parcial', main="")
```

```

win.graph(width=8,height=5,pointsize=12)
boxplot(x1 ~ cycle(x1),xlab='meses',col='grey',ylab='Precipitação PVH ')
adf.test(x1)
win.graph(width=8,height=5,pointsize=12)
y=diff(x1,lag=12,differences = 1)
layout(1:2)
m = acf(y, plot=F); m$lag = m$lag*12; plot(m, xlab='Defasagem',ylab='Autocorrelação',
main=")
m1 = pacf(y, plot=F); m1$lag = m1$lag*12; plot(m1, xlab='Defasagem',ylab='Autocorrelação
Parcial', main=")
## testando o modelo
fit <- arima(x1, order = c(1,0,2), seas = list(order = c(1,0,1), 12))
fit
t( confint(fit) )
fit$coef
win.graph(width=8,height=5,pointsize=12)
layout(1:2)
m = acf(fit$residuals, plot=F); m$lag = m$lag*12;plot(m)
m1 = pacf(fit$residuals, plot=F); m1$lag = m1$lag*12; plot(m1)
win.graph(width=8,height=5,pointsize=12)
tsdiag(fit)
z=fit$residuals
win.graph(width=8,height=5,pointsize=12)
qqnorm(z)
qqline(z)
## teste para normalidade
shapiro.test(z)
## Teste de Ljung-Box e Box-Pierce até 24 lags
Box.test(fit$residuals, lag=24, type ="Ljung-Box",fitdf=2)
Box.test(fit$residuals, lag=24, type = "Box-Pierce", fitdf=2)
## Teste kolmogorov smirnof
ks.test(fit$residuals,"pnorm",mean(fit$residuals),sd(fit$residuals))
## Analise das previsões
win.graph(width=8,height=5,pointsize=12)
prev <- predict(fit, n.ahead = 12)
ts.plot( cbind( x,(predict(fit,12)$pred)),xlab='Ano',ylab='precipitação',main="",col='blue',
bty="l", lty = 1:2 )
prev

##script Precipitação VHA
library(tseries)#ativa o pacote
library(TSA)#ativa o pacote
getwd()
setwd("C:/Users/Cliente/Downloads/Nova pasta/dados e script TCC")
getwd()
x <- read.table("txt.prc.vh.txt", header=T)
x
x=ts(x,frequency=12,start=c(2008,8))
x1 = window(x, start=c(2008, 8), end=c(2015, 8))
class(x); start(x); end(x); frequency(x)

```

```

class(x1); start(x1); end(x1); frequency(x1)
win.graph(width=8,height=5,points=12)
plot(x1,xlab='Ano',ylab='Precipitação VHA',main='',col='green', type='o', bty="l")
win.graph(width=8,height=5,points=12)
plot(decompose(x1))
## Teste Cox-Stuart - verificar a tendência da série
cox.stuart=function(x, trend=c("+","-")){
  trend=match.arg(trend)
  N=length(x)
  p1=x[1:(N/2)]
  p2=x[(N/2+1):N]
  sinais=sign(p2-p1)
  if(trend=="-") T2=sum(sinais<0) else T2=sum(sinais>0)
  n=length(sinais[sinais!=0])
  if(trend=="+") tcrit=qbinom(.95,n,.5) else tcrit=qbinom(.05,n, .5)
  if(trend=="+") pv=1-pbinom(T2, n,.5) else pv=pbinom(T2, n, .5)
  if(round(pv, 4)==0) pv="<0.0001" else pv=round(pv, 4)
  data.frame("N"=N, "n"=n, "estatistica"=T2, "valor.critico"=tcrit, "tendencia"=trend, "p-
valor"=pv)}
cox.stuart(x1,trend="+")
##Teste Fisher - verificar a sazonalidade da série
win.graph(width=8,height=5,points=10)
pgram=spectrum(x1, fast=FALSE, detrend=FALSE, log="no", plot=T, main="Periodograma
da série")$spec
n=length(pgram)
pvalue.g=function(g.obs, n) n*(1-g.obs)^(n-1)
g.obs=max(pgram)/sum(pgram)
pvalue.g(g.obs,n)
win.graph(width=8,height=5,points=12)
layout(1:2)
m = acf(x1, plot=F); m$lag = m$lag*12; plot(m, xlab='Defasagem',ylab='Autocorrelação',
main='')
m1 = pacf(x1, plot=F); m1$lag = m1$lag*12; plot(m1,
xlab='Defasagem',ylab='Autocorrelação Parcial', main='')
win.graph(width=8,height=5,points=12)
boxplot(x1 ~ cycle(x1),xlab='meses',col='grey',ylab='Precipitação VHA')
adf.test(x1)
win.graph(width=8,height=5,points=12)
y=diff(x1,lag=12)
layout(1:2)
m = acf(y, plot=F); m$lag = m$lag*12; plot(m, xlab='Defasagem',ylab='Autocorrelação',
main='')
m1 = pacf(y, plot=F); m1$lag = m1$lag*12; plot(m1, xlab='Defasagem',ylab='Autocorrelação
Parcial', main='')
## testando o modelo
fit <- arima(x1, order = c(1,0,1), seas = list(order = c(1,0,1), 12))
fit
t( confint(fit) )
win.graph(width=8,height=5,points=12)
layout(1:2)

```

```

m = acf(fit$residuals, plot=F); m$lag = m$lag*12; plot(m)
m1 = pacf(fit$residuals, plot=F); m1$lag = m1$lag*12; plot(m1)
win.graph(width=8,height=5,pointsize=12)
tsdiag(fit)
z=fit$residuals
win.graph(width=8,height=5,pointsize=12)
qqnorm(z)
qqline(z)
## teste para normalidade
shapiro.test(z)
## Teste de Ljung-Box e Box-Pierce até 24 lags
Box.test(fit$residuals, lag=24, type = "Ljung-Box", fitdf=2)
Box.test(fit$residuals, lag=24, type = "Box-Pierce", fitdf=2)
## Teste kolmogorov smirnof
ks.test(fit$residuals,"pnorm",mean(fit$residuals),sd(fit$residuals))
## Analise das previsões
win.graph(width=8,height=5,pointsize=12)
prev <- predict(fit, n.ahead = 12)
ts.plot( cbind( x,(predict(fit,12)$pred)),xlab='Ano',ylab='precipitação',main="",col='blue',
bty="l", lty = 1:2 )
prev

```